NONSEPARABLE UTILITY AND LABOR SUPPLY.

by

Finn E. Kydland

Hoover Institution
Stanford University

May 1983
Revised August 1983

Prepared for the NBER Conference on Macroeconomics, Cambridge, Mass., July 7 and 8, 1983. This research was supported by the National Science Foundation. The author has been on leave from Carnegie-Mellon University.
1. Introduction

In the past ten years, there has been a revival of interest in equilibrium models of the business cycle. The view had been that accounting for the persistence of cyclical movements in output and other important aggregates would be an extremely challenging, if not impossible, task with such models. The important paper by Lucas (1972) led to models which are consistent with monetary shocks resulting in persistent equilibrium movements of real aggregates. The key to these results was the richness of the information structure, and it was only natural that less effort went into developing the details of propagation mechanisms that might be present on the real side of the economy. Although there was an underlying notion that the economy is inhabited by optimizing agents who process information efficiently within the specified information sets, it was not always necessary to work out the optimization part explicitly in order to bring home the point of this research. Instead, simple formulations emphasizing intertemporal substitution effects and intended to mimic such dynamic behavior were often used.

Several factors may have contributed to a recent trend in the direction of analyzing models of the aggregate economy in which decisions are derived from explicitly formulated optimization problems. First, there was the realization that the parameters of these optimization problems are part of what we will ultimately want to estimate in order to evaluate systematic tax changes and other government policies with some confidence. Some people realized that, in addition to monetary shocks, technological and other real shocks might also be important factors in the business cycle. In order to analyze the effects of such shocks, the real propagation mechanisms deriving from preferences and technology became more essential. It is clear,
however, that the importance of such mechanisms does not depend on the nature of the shocks. Their formal inclusion in monetary theories would presumably also make these theories more powerful in explaining the observations.

Finally, the attractiveness of optimizing, competitive equilibrium theories largely stems from the discipline they impose on the researcher wishing to understand the nature of business cycles. Such models allow for considerable richness in dynamic aggregate behavior while keeping the number of free parameters small. Recent methodological advances have made their analysis feasible. Rational expectations is part of the equilibrium definition and means that individual agents understand the structure of the economy in forming their expectations of future prices which are important determinants of current behavior. In other words, the structure used for this purpose by economic agents coincides with the one that results from their aggregate behavior. Such models have to abstract from many things in order to serve their purpose and can obviously not yet be expected to fit the data very well according to usual goodness-of-fit criteria. Adding parameters that might be motivated by for example disequilibrium phenomena would improve the fit, and some basis in realism could probably be given. As with most scientific effort, however, it is generally the case that insisting on too much realism reduces the chances of learning anything useful.

If one accepts this basic approach, it is obviously still the case that a great variety of models would fit that general description. That is, we have to find out what are the right model elements to put into such a theory. Given the novelty of this line of research, however, it is only natural that we have not gotten very far in determining what they are.
In this paper, I shall concentrate on one particular model element that has been proposed. It is a utility function which is intertemporally non-separable in leisure. The next section reviews some of the evidence so far. Section 3 looks into some of the characteristics of behavior that results from nonseparable utility. It is shown in Section 4 that an inherently time-separable utility function combined with a home production function which includes a durable affected by nonmarket activity yields behavior similar to that implied by the nonseparable function. In Section 5, I argue that the period length (e.g., yearly versus quarterly) may make a substantial difference in how to interpret or compare empirical results based on different data sets. The possibility of gaining evidence from panel data is discussed in Section 6, while Section 7 suggests some policy implications. Concluding comments are offered in the last section.

2. Evidence from Aggregate Data

The idea that intertemporal substitution of leisure is important in determining employment fluctuation over the business cycle was expressed very clearly in the empirical paper by Lucas and Rapping (1969). They in fact suggested that a nonseparable utility function was needed to explain the observed movements of employment relative to wage rates in the aggregate economy. Using yearly data, they found that long-run labor supply was inelastic, while the short-run elasticity was 2.2 for their main specification. Altonji (1982) found, however, that "after corrections for minor data errors, their work indicates that the elasticity of labour supply with respect to temporary wage changes is around 4.6." This is in contrast with the much lower elasticities from Altonji's own empirical specifications as
well as with those of for example Hall (1980) and Clark and Summers (1982). Hall's results are hard to interpret in this context, however, due to his combining wage and interest rate effects into a single variable and not accounting for expected future wages. For those reasons, Altonji argues that Hall's estimates are likely to be biased downwards.

It is good to be clear on how to interpret the short-run elasticity obtained by Lucas and Rapping. The terms involving wage rates in their labor supply function can be written as $\beta_1 \ln(w_t^*/w_t^*) + (\beta_1 - \beta_2) \ln(w_t^*)$, where $w_t^*$ can be interpreted as the permanent or normal real wage, and the first term represents the percentage transitory difference from that wage. Thus, $\beta_1$ is the short-run labor supply elasticity. In their main empirical specification, the expected permanent wage is updated as follows:

$$\ln(w_t^*) = \lambda \ln(w_{t-1}) + (1 - \lambda) \ln(w_{t-1}^*) + \text{trend}. $$

Thus, if the wage has a permanent and a transitory component, but only the sum can be observed, then $\lambda$ can be interpreted as the variance of the permanent wage shock relative to the sum of the two variances. A Koyck transformation using this relation puts the supply function on a form involving current and lagged wage rates and lagged manhours. Lucas and Rapping report a short-run elasticity of 1.4, which is the coefficient for the current wage rate in this transformed equation, but this coefficient now involves $\beta_2$ and $\lambda$ as well as $\beta_1$. The implicit value for $\beta_1$ is 2.2 and would represent the relevant measure of short-run labor-supply response to, say, a tax change which is known to be transitory. For some policy issues, of course, the compensated elasticity is the appropriate one to use.
Lucas and Rapping pointed out the potential shortcomings of basing the forecasts on such a simple wage process in an aggregate model, but suggested that it might not be too unreasonable for the period considered, which was 1929-65. Altonji reestimated various versions of their basic model using data up until 1981. He, however, first estimated reduced-form equations for the wage rate and the price level from which expectations of their future values were determined. These equations included the variables that entered the labor supply and demand equations, but consistency between the wage relation and the resulting labor market equilibrium was not imposed. Also, variables such as the price level, the nominal interest rate, and real output were assumed to be exogenous with respect to the employment and wage determination. Thus, although the relations describing the formation of expectations here included a lot more variables, expectations could still not be considered rational in the sense described in the introduction, and one should perhaps not put too much weight on the fact that Altonji even obtained negative labor supply elasticities for many of the model specifications he considered. Similar comments can be made with regard to the expectations used in Clark and Summers (1982).

In a pair of papers, Kydland and Prescott (1980, 1982) incorporated the assumption of intertemporally nonseparable utility in general equilibrium models of the business cycle. We shall here consider the most recent paper. The basic approach is to construct abstractions of the aggregate economy and compare their variance-covariance properties with those of the U.S. data for the postwar period. The models were constructed so as to facilitate the imposition of a priori knowledge while leaving just a few parameters over which to search for a good fit. Methods were used that were thought to be
robust to things like measurement errors which do not come out of the theory, timing conventions, and other elements that are part of the maintained hypothesis, and also to changes in the government policy environment (or agents' expectations thereof).

The model specifically combined the assumption of non-time-separable utility in leisure with a time-to-build investment technology, and also an information structure related to the shocks which were assumed to be largely technological. The specification of the utility function was

$$\sum_{t=0}^{\infty} \beta^{t} u[c_t, a(L)l_t],$$

where $c_t$ is consumption in period $t$, $l_t$ is leisure, $\beta$ equals $1/(1+\rho)$ where $\rho$ is the rate of time preference, $L$ is the lag operator, and $a(L) = \sum_{i=0}^{\infty} a_i L^i$. The form of this lag distribution describes the importance of past leisure choices in current utility. Several such distributions are conceivable. Without loss of generality, the weights can be made to sum to one. For empirical purposes, the distribution should be parsimoniously characterized in terms of a small number of parameters. For example, one might cut off the distribution after a couple of lags and let the weights otherwise be free. The case of one lag is the specification used in Sargent (1979). Kydland and Prescott (1982) chose the alternative of a geometrically declining weight distribution whose shape was determined by a depreciation rate, but with the weight on current leisure a free parameter. This allows for weights going back more than a couple of quarters while keeping the model computationally tractable. The weight distribution is then characterized by two parameters, the weight $a_0$ on current leisure and the depreciation
rate $\eta$ for the weights on past leisure choices. The standard separable specification of course corresponds to the special case of $\alpha_0$ being equal to one.

The utility function $u$ is of the constant-relative-risk-aversion variety, namely $c_t^u \left[ u(L_t) \right]^{1-\gamma} / \gamma$, $0 < u < 1$, $\gamma < 1$, whose relative degree of risk aversion is $1-\gamma$. The limit as $\gamma$ approaches zero is the logarithmic function. The parameter $\mu$ is close to the long-run fraction of total time allocation spent in market activity.

Given its simplicity, this model matched the data surprisingly well. The key model elements were compared with their most prominent alternatives. Of particular interest here is the comparison between the nonseparable utility function and the standard time-separable one. Postwar U.S. data indicate that the percentage fluctuation of cyclical employment from its growth path has been substantially larger than for cyclical productivity.\(^6\) This was also the case in the model version that included nonseparable utility, although not by as much as in the data.\(^7\) With the separable specification, the percentage fluctuation of employment was substantially less than for productivity. It is interesting to note, however, that the performance along this dimension depends on the rest of the model, in particular on the assumption about investment technology. While the nonseparable preference formulation always led to substantially more employment fluctuation than the separable one did, the percentage employment fluctuation was still much smaller than that of productivity when a technology with moderate cost of adjustment was employed instead of the time-to-build assumption. This change is presumably mainly due to differences in interest rate effects that arise in equilibrium for the two specifications of investment technology.
The results provide guidance in terms of where the greatest potential is for improvement on the model. The comparison of the model with the data was in terms of a set of statistics. The sample data of course represent one particular draw. Imagine now that repeated draws of the same length (in this case 118 quarters) are made from the model. Then standard deviations for each of the statistics can be computed and used as measures of how well the model performed along each of the individual dimensions as opposed to in a global sense. For example, the standard deviation of cyclical employment in the model was almost a third lower than that of the data. Also, there was the related problem that the model's correlation of productivity with output was too high. Both of these problems could easily be due to measurement errors. The abstraction had only one type of labor skill, and hours of employment must therefore be interpreted as being measured in efficiency units. The data, of course, are not in such units. During booms, a relatively larger proportion of less skilled workers (e.g., youths) enter the labor force. This suggests a cyclical overmeasurement in the employment data during booms and a similar undermeasurement during downturns. Thus, there is a procyclical measurement error relative to the theory. In addition, employment statistics are probably quite noisy in the first place, which indicates there may also be an error component which is nearly independent over time.

As suggested above, one dimension on which the model can be judged is in terms of the ratio of cyclical employment fluctuation to that of productivity (each measured by standard deviations of the cyclical components relative to their levels). Postwar U.S. data indicate that this ratio is about 1.7. If employment is overmeasured by, say, 25 percent during booms
and undermeasured by the same percentage during recessions, then this ratio in the model increases from 1.20 to 1.81, while the correlation between productivity and output drops from 0.90 to 0.75. If, in addition, the standard deviation of an independent error component were 0.5 percent of the average level of employment, then the above-mentioned ratio would be 1.67, and the correlation between productivity and output would show a further drop to 0.61 as compared with 0.34 in the data. Including such measurement errors would also improve the performance of the separable model, but the relative differences in properties would still be present.

Theory generally does not tell us what are reasonable assumptions to make about measurement errors. While the above assumptions appear realistic and would have improved the closeness of fit, one should not stop there, but rather search for ways of improving the model without relying on these errors.

There is always the possibility that even if a model element works well, somebody might present a different hypothesis that also explains the data well. Thus, various forms of cross-verification are needed before one can be confident that a particular element, in this case the intertemporally non-separable utility function, should be part of future business cycle models. Subsequently, we shall discuss the possibility of gaining evidence of non-separability from panel data. There are also other kinds of data that go into considerable detail about various aspects of labor markets. For example, one could investigate whether the assumption of nonseparable utility and the resulting possibility that the reservation wage is countercyclical helps us understand relative movements of wages and employment in unionized versus nonunionized sectors over the business cycle. Such cyclical movements are analyzed in McDonald and Solow (1982). Another possibility is to see if the
failure of the alternative wage rate to have the influence predicted from the
theory in empirical tests of the efficiency of employment contracts as in
Ashenfelter and Brown (1982) can be understood better in this light.

3. **Implications of Nonseparable Utility**

It may be helpful to give a further indication of the differences in
behavior implied by the two competing utility specifications. As already
pointed out, within a general equilibrium framework there may be several
more or less simultaneous effects as the economy is being hit by shocks and
the wage rate as well as other prices find their equilibrium. Taking an
opposite extreme of assuming no interest rate effects and an exogenous wage
process will illustrate the differences more clearly. The relevance for
panel data should not be taken too seriously, however. There are many kinds
of experiments that can be imagined for different stages of the life cycle.
Instead, I shall assume a sufficiently long-lived individual (interpreted as
a representative household in an aggregate context) that his behavior is
given by the same decision rule in every period. He is assumed to live in a
stationary environment and will have zero long-run nonhuman wealth.

If one only makes the assumptions that are standard in studies of qual-
itative properties of behavior arising from different utility functions, it is
clear that the separable function is consistent with a great variety of behav-
ior. Barro and King (1982) demonstrated, among other things, that it can in
general be consistent with large percentage employment fluctuation associated
with relatively small wage fluctuation. This result is of little help in this
context, however. If we consider, for instance, the constant-relative-risk-
aversion function, it is clear that we are not allowed a great deal of freedom
in choosing parameter values if we want it to be consistent with the observations. Labor economists have found that the average long-run fraction of nonsleeping time spent in market activity is close to one-third. This suggests that the parameter $\mu$ cannot be far from this value. Also, studies of magnitudes of risk premia suggest that the utility function should not be too close to linear. The degree of risk aversion associated with the logarithmic function is probably a lower bound for what can be considered realistic.

As an illustration, consider the utility function described in Section 2. The per-period budget constraint will be

$$A_{t+1} = (1+r)(A_t + w_t n_t - c_t),$$

where $A_t$ is net real assets at the beginning of period $t$, $r$ is the interest rate, $n_t$ is the time allocated to market activity, $w_t$ is the real wage rate, and $c_t$ is consumption in period $t$. The interest rate is assumed here to be a constant and equal to the rate of time preference $\rho$. Long-run net assets are zero. The wage rate can be thought of as having a permanent and a transitory component. With the dependence of utility on current and past leisure choices characterized by the two parameters $\alpha_0$, the weight on current leisure, and $n$, the depreciation rate for past leisure choices, it can be shown that the long-run value of $n$ must satisfy the condition

$$n = \mu/\left(\mu + (1-\mu)\left[\alpha_0 + (1-\alpha_0)n/(n+\rho)\right]\right)$$

independently of the long-run real wage $w$. For the special case of time-separable utility, we have $n = \mu$. In general, if the parameter $n$ is substantially larger than the rate of time preference $\rho$, and $\alpha_0$ is not extremely small, both of which are reasonable assumptions, then $n$ is very close to $\mu$. 
If we choose a very small value for \( \mu \), then the short-run labor supply elasticity can be quite high even with time-separable utility. It is also the case that the closer to one \( \gamma \) is chosen, the larger is this elasticity. Such values of \( \mu \) and \( \gamma \) would be inconsistent, however, with a priori knowledge. If one maintains the quantitative discipline of restricting the utility function so as to be consistent with long-run relations that can be found in the data, or with properties that have been established elsewhere in the literature, then an elasticity of 1.5 is about the most one can get.

In the general case, the parameters describing the lag distribution for leisure in current utility are of course what we are ultimately interested in determining. For the present purpose, a representative illustration might be the combination of \( \alpha_0 = 0.6 \) and \( \eta = 0.5 \). Remembering that the weights are assumed to sum to one, this means that there is a weight of 0.6 on \( l_t \) in current utility, with weights 0.2 on \( l_{t-1} \), 0.1 on \( l_{t-2} \), and so on. The remaining parameters are taken to be \( \mu = 0.33 \), \( \gamma = -1 \), and \( \rho = \sigma = 0.015 \) on a quarterly basis.

It is interesting to compare the variance-covariance properties of this model with the special case of time-separable utility. What is reported as the short-run labor supply elasticity is the coefficient of the labor supply decision rule with respect to the transitory wage rate and evaluated as an elasticity at the long-run level. For the time-separable case it is 1.3, while the nonseparable formulation yielded 3.6. In both cases, consumption fluctuated relatively less than the transitory wage, but with about 60 percent larger average amplitude for the nonseparable than for the separable case. As already mentioned, however, the magnitudes of such differences,
say as reflected in relative fluctuations of cyclical employment versus cyclical real wage, do not carry over to a general equilibrium model such as the one employed by Kydland and Prescott (1982). While the metric for evaluating such differences in that context may not be all that clear, it does seem that capital accumulation and interest rate effects reduce them.

There were also large differences between the two models in terms of lagged correlations. The first-order autocorrelation for hours worked was -0.28 for the nonseparable case and -0.07 for the separable one. The correlations between hours worked and last period's transitory wage were similarly -0.36 versus -0.08. For both models there is a slight wealth effect in periods subsequent to the one in which the transitory wage change takes place. In the nonseparable case, there is an additional effect through the influence of past leisure choices on the marginal rate of substitution between leisure in the current and next period.

The properties of the nonseparable example above are quite robust to changes in $\alpha_0$ and $\eta$ within a broad range. Of course, when $\alpha_0$ gets close to one, the properties become similar to those of the separable specification irrespective of $\eta$. In general, the value of $\alpha_0$ to a large extent determines the short-run labor supply elasticity which changes little for different values of $\eta$. The lagged correlations, on the other hand, are more sensitive to $\eta$. Relative to the example above, an increase in $\eta$ to 0.7 results in a first-order autocorrelation of hours worked of -0.39 and in a correlation between current hours and lagged wage of -0.47. These results suggest things to look for in the data at an informal level as preliminary evidence of significant non-time-separability.
If the wage process has both a transitory and a permanent component, then it is important to correct for the permanent part in estimating the short-run elasticity. Failure to do so, say in regressions of hours on wage rates, will bias the estimates of the elasticity downwards to an extent that depends on the relative variances of the two components. Then there is the problem that only the sum of the permanent and transitory components is observed, which is an issue that Lucas and Rapping gave considerable attention. Such problems may very well have contributed to the rather low short-run elasticities reported in some studies.

The results of this section suggest that if one estimates the parameters of a utility function which is specified as a time-separable one, then there is implicitly a limit on the magnitude of short-run labor supply elasticity that one can expect to find in the data. Of course, if one leaves free some of the parameters that we took more or less as fixed from other considerations, and individuals really behave as though their utility is nonseparable, then the values of these parameters might get moved around in the estimation, perhaps so as to make the elasticity look larger.

4. An Underlying Household Production Theory

What we are interested in finding out is whether the abstraction of a utility function which is intertemporally nonseparable in leisure as in the preceding section "works" in the sense of being consistent with certain regularities in the data which we wish to understand. In this section, I suggest a more basic foundation for this model element, which to some people may make it more plausible. A potentially important by-product might be insights into the nature of weights on past leisure choices that
we can expect to see. Of course, if the period length is sufficiently short, there is probably little disagreement about the possibility that past commodity choices can affect instantaneous utility. The enjoyment of a dinner at a three-star restaurant is affected by whether one had a big lunch that day. Long hours of market activity result in fatigue. The data we are trying to explain, however, are typically quarterly or even yearly.

The home production theory shows that a nonseparable utility function of the form above can be viewed as a stand-in for a situation in which non-market activity, among other things, either adds to or maintains a possibly unobservable durable which itself affects preferences. One can imagine several possibilities. The utility of services provided by durables in the form of market goods (such as the home with related durables, cars, etc.) may depend on time input. In some cases, hiring somebody else's time can be an alternative to some forms of one's own time input. In other cases, as in the production of "high quality" children, parents have family-specific abilities that make them not easily substitutable. Once the decision to have a child has been made, there is presumably an intention or even commitment to spending a fair amount of time with them, although some degree of lumpiness of that time is acceptable. If both parents work hard for a while in response to a temporarily high wage, then the marginal product of at least one of them spending more time in the home rises. Other, generally less observable, durables (e.g., health) may be important as well.

We shall consider the simplest case and assume that a portion of non-market time contributes to the accumulation of a home durable whose stock at time t we shall denote by $d_t$. Thus, total time (or its nonsleeping portion) can be allocated as follows:
\[ T = n_t + \lambda_{1t} + \lambda_{2t}, \]

where \( n_t \) is market activity, and \( \lambda_{2t} \) results in accumulation of the durable as follows:

\[ d_{t+1} = (1-\delta)d_t + \lambda_{2t}. \]

Current utility is a function of consumption of market goods \( c_t \), "pure" leisure \( \lambda_{1t} \), and (the services of) \( d_t \). The behavioral implications of the nonseparable utility function of the preceding section are equivalent to those of the special case of the present structure in which a fixed proportion of nonmarket activity is spent on producing the durable, and \( \lambda_{1t} \) and \( d_t \) are perfectly substitutable in preferences. Thus, the role of the assumption that the weight on current leisure, \( \alpha_0 \), is free and generally substantially larger than the other weights is to allow for the presence of leisure in the form of \( \lambda_1 \) in addition to the leisure stock.

An alternative specification would be the opposite extreme on the production side in the sense of letting the two forms of nonmarket activity be perfectly substitutable, but with less substitutability in preferences. For example, with the form \([c_t^{\mu_1} \lambda_{1t}^{\mu_2} d_t^{(1-\mu_1-\mu_2)}]^{\gamma/\gamma} \) and the same budget constraint as in the previous section, long-run \( n \) is

\[ n = \frac{\rho + \delta}{(\mu_1 + \mu_2)\rho + \delta} \mu_1. \]

We see again that if \( \delta \) is substantially larger than \( \rho \) and the weight on past leisure choices through \( d_t \) is not too large, then the ratio is close to one, and long-run labor supply is close to the exponent \( \lambda_1 \) on consumption in the utility function. For the special case of no durable, that is, \( \mu_1 + \mu_2 = 1 \),
long-run $n$ is equal to $u$. Thus, the steady state properties can be made to be equivalent to those of the nonseparable function of the previous section, but the dynamic properties will not be exactly alike.

While there is at least one formulation of home production and preferences which is behaviorally equivalent to a utility function in which the intertemporal nonseparability is characterized by our two parameters, this discussion also makes it clear that, in general, that particular form must be viewed as an approximation. For some home production formulations, the behavior is going to be as if the weights in the nonseparable utility function vary slightly over time. The general declining pattern of the weights will be maintained, however, and, even as a basis for understanding individual behavior, the parsimonious two-parameter formulation may still be a reasonable approximation. This may be the case also if market goods are an input along with time in the production of the durable.

One might argue that if this is the story, then it is preferable to use this more basic theory directly. An answer is that, at least at the aggregate level, there are probably no observations that would be of any help, and therefore one may as well base empirical work on the nonseparable utility function.

The general idea of intertemporally nonseparable utility has been considered by several economists in the context of consumption of market goods. An example is the assumption of habit formation, although in that case, contrary to our assumption about leisure, goods in neighboring periods are complements. It is not clear to me to what extent this idea arose from a demonstrated need for it in the data. If significant evidence of the importance of lagged consumption is found, an alternative interpretation is that
the apparent effect is due to the omission of consumer durables in the analysis and demonstrates the need for including durables explicitly in the model. In this case, there are indeed data on aggregate durable consumption expenditures.  

5. **Length of Observation Period**

Some empirical studies use data that are available on a yearly basis only. This is the case in studies using the Michigan Panel Study of Income Dynamics, and also in some studies of aggregate behavior such as Lucas and Rapping (1969). A potentially serious issue for this type of model is time aggregation. If for example one looks at the impulse response function for labor supply associated with a positive transitory wage change in period $t$, there is a relatively large contemporaneous effect on employment. In period $t+1$, however, there is a sizable negative effect, and then the response function moves steadily back towards the long-run level from below, although for large depreciation rates it would generally go above the long-run level in period $t+2$ and oscillate from then on. It is reasonable to think that when four and four quarters are added or averaged to yield yearly observations, then behavior on the basis of a nonseparable utility function would result in data that look closer to the observations that would have been generated from a time-separable function.

As an illustration, consider the basic example of Section 3. If we regress hours of work on the temporary wage rate and evaluate the coefficient as an elasticity at the long-run level, we get numbers close to the short-run labor supply elasticities reported there. If we temporally aggregate quarterly observations into yearly ones, the elasticity obtained from this
regression for the nonseparable model drops from 3.7 to 2.6. The first-order autocorrelation for hours worked and the correlation between hours in period t and the wage rate in period t-1 are both -0.23. Thus, while all three numbers have declined substantially, there is still a significant difference between these figures and the corresponding ones for the time-separable model.

Some people have argued that there are severe restrictions on the extent to which individuals can adjust their hours or move in and out of the labor force due to contracts, implicit commitments, adjustment costs, or for other reasons. Such arguments are likely to work in the opposite direction, that is, play a bigger role the shorter the time period is. If they are valid, then certainly the month and perhaps even the quarter would be too short to get the intertemporal substitution effects that result from the nonseparable utility function. If the nature of the durable entering the home production function as described in Section 4 is such that half-a-year or a year is a reasonable period length in our utility function, then it is not inconceivable that monthly data could still look as if generated by a cost-of-adjustment story and thus presumably result in a negative estimate for the weight on lagged leisure.

It is well known that much of economic data, at least at the aggregate level, show a great deal of persistence in the sense that once series move significantly to one side of the growth paths, they usually take several quarters to move back. While there are persistent employment effects from temporary wage changes with nonseparable utility, they do not have the pattern just described. This suggests that if there is nothing else in the model that generates persistence and especially if the estimates are sensitive to the properties assumed for the wage process, then one has to be
careful not to put burden on the nonseparable utility function to explain such persistence which then might result in negative weights on lagged leisure. For some models, a careful choice of instruments will clearly be essential.

6. The Use of Panel Data

The Michigan Panel Study of Income Dynamics now covers up to 15 years starting in 1967 for a large number of individuals. For each person, there are data on annual hours of market work, average hourly wage rate for the year, household expenditures on food in the home as well as on meals eaten out, age, race, education, post-school training, number of children divided into two age groups, and many other things for both males and females. Given the way in which the data on some variables were obtained, however, they are likely not to be very accurate. The potentially most severe measurement problem for our purpose is the following. Each individual was asked to report total wage income for the previous year along with the number of hours worked. The wage rate was then computed by simply dividing the two numbers. Aside from potential problems with using an average wage, there is the problem that if hours are reported with error, then there will obviously be a highly correlated error in the wage rate in the opposite direction. The directions of these errors are such that, whatever evidence there is for nonseparability in the true values, the actual data will show less evidence. Attempts at getting a feel for the likely magnitude of this effect would be valuable.

Given such data problems and likely sensitivity of the results to other parts of the empirical specification, it certainly seems worthwhile to look
at the data in several different ways. Considering the differences in covariance properties found above for the two utility specifications, an informal investigation of the properties of the data in relation to theory might be useful. A crucial part of such an effort would be to filter out from the wage series what can be thought of as the permanent wage component or the individual's conditional expectation thereof since it cannot be directly observed. A simple trend would probably not do it. The average reported wage rate deflated by the consumer price index grew quite rapidly up until 1973, at which time it pretty much leveled off for the rest of the sample. Eventually, this decline in wage growth probably was considered as rather permanent, although it was harder to distinguish from a temporary decline back in 1974. Rather than using a deterministic trend, a stochastic one based on Kalman filtering as in Crawford (1979) may be more appropriate.

Aside from the types of measurement problems mentioned above, there is the question of whether, in principle, the wage rate in the panel data is the appropriate measure of compensation for our purpose. The slowdown of wage growth in the mid-seventies may be partly due to workers receiving a larger portion of their salaries in the form of fringe benefits. The biennial U.S. Chamber of Commerce survey of large manufacturing establishments, for instance, shows that total fringe benefits as a percent of payroll have increased from 20.3 percent in 1957 to 37.3 in 1977. Much of this increase occurred between 1969 and 1975 when fringe benefits rose from 27.0 to 36.1 percent.\footnote{19} On the other hand, increased rates of inflation have probably pushed most workers into higher marginal tax brackets.\footnote{20} These two effects may to some extent cancel. The former may indeed partly be a response to the latter. What we know about compensation rules for workers covered by
collective bargaining also suggests difficulties in interpreting the reported wage-hours combination in a given time period for many workers.  

In panel data, life-cycle considerations are important. As an illustration, I looked at a subsample of all the males who worked in every period. This subsample was divided in two by age. The average total real-wage increase over the 12-year period was 35 percent for the young workers and 15 percent for the old. As one would expect from work by Ghez and Becker (1975) and others, there is a significant human capital component to the wage increase that is particularly important for the young. One should also consider to what extent the nature of the intertemporal allocation of leisure in response to wage rates interacts with the investment decision in early years when most of the human capital accumulation in the form of on-the-job training takes place. By the same token, one has to be careful about the members of the sample who are near or past the age of 60 and thus near retirement at the end of the sample period.

The female subsample may be particularly important for our purpose. With the household as the relevant decision unit, much of the intertemporal allocation of time is likely to have taken the form of the wife adjusting her hours or moving in and out of the labor force. For the household as a whole, however, it is also important to consider the possibility that dropping out of the labor force by one member for a period of time may be partly compensated for by the other member working more in the same period.

Unfortunately, it appears to be a particularly challenging task to get information on these issues out of the subsample on wives. This sample is relatively small compared to the rest of the panel data set. If we use the first 11 years and include only those who were continually married and were
between the ages of 30 and 63 throughout this sample period, we are left with 399 individuals.\textsuperscript{22} Out of these, 81 did not report working in any of the 11 years. If we include only those who worked at least six years, say, this would limit the sample to 218 females. Then there are quite a few who suddenly one year report an unusually low number of hours and a correspondingly big increase in the wage rate (or the two changes could be in the opposite directions), say, by a factor of eight or ten, but then revert to normal figures the year after. It is hard to see how these examples could be anything other than extreme cases of the measurement problem I mentioned above. If we exclude the individuals for whom there was a real wage change from one year to the next by more than a factor of three, there will be 156 wives left. A factor of four would make it 178, and a factor of two 119. In either case, we would be down to a rather small number of individuals, with only the extreme cases of measurement errors removed. A challenging and potentially important project is to find ways to confidently use much of the data that would otherwise be lost through considerations such as the above.

The model in Section 3 indicates some hope that at least the less extreme cases of measurement errors can be dealt with in testing for non-separability. For example, if we add serially uncorrelated measurement errors to the hours (or equivalently to the real wage) but for simplicity not to gross wage income, the autocorrelation for measured hours as well as the correlation between hours and the lagged wage rate change little, even though the measure of short-run supply elasticity might decline substantially. This is the case for moderate measurement errors, say with standard deviation of up to half the standard deviation of the transitory
wage shock. These correlations were among the most distinguishing features characterizing behavior resulting from nonseparable utility. On the other hand, even without measurement errors, the standard deviations of these statistics under repeated draws of, say, 12 periods from the model are quite high. This is another piece of evidence that avoiding too much reduction in the sample could be quite valuable.

Finally, to the extent that the home production story associated with children is viewed as a promising one, this suggests that the degree of intertemporal substitution may be different depending on the number of children. The panel data may be particularly well suited for investigating that possibility.

7. Policy Implications

As was pointed out in Kydland and Prescott (1980), intertemporal substitution effects have an important bearing on some public finance issues. Assume for instance that public goods enter individuals' utility, but in such a way that their tastes for these goods change stochastically (larger demand during wars, for example). The role of the government is to provide these goods. For simplicity, it can finance public expenditures in either of two ways: by changing taxes on labor income or by issuing government debt. One can show that deadweight loss is minimized by choosing a policy which involves less change over time in the tax rate on labor income the more intertemporally substitutable leisure is, or, alternatively, the more significant the nonseparability of utility is (put somewhat imprecisely). Rather, most of the fluctuation in demand for public goods should be picked up by letting the public debt fluctuate. While this is an issue also when
preferences are time-separable, it is quantitatively more important when utility is intertemporally nonseparable.

Government policies under which the policy instrument remains constant (or nearly so) have been advocated in different contexts at least going back to Friedman (1948). His reasons were different, however. His main argument was that we still do not have the requisite understanding of the economy, and that timing and magnitudes of effects can be long and variable. Policies may therefore have destabilizing rather than stabilizing effects. More recently, it was pointed out that dynamic optimal tax policy is not time consistent and therefore not credible. A reasonable policy, one can therefore argue, is one which is simple and well understood and where a change would be quite transparent. The discussion above potentially reinforces those arguments. That is, if leisure is intertemporally highly substitutable, then, even if we did have the requisite understanding of the economy and the present government could credibly commit itself and future governments, a policy of keeping the tax rate nearly constant would still be the best policy.

8. Conclusion

In this paper, I have entertained the possibility that a utility function which is intertemporally nonseparable in leisure is an important model element in an equilibrium theory of the business cycle based on optimizing behavior. The evidence so far is mixed. Some of the empirical studies in which labor supply elasticities have been estimated can be criticized because of assumptions made with regard to expectations. An example of a consistent general equilibrium model is in Kydland and Prescott (1982). One can argue, however, that also in that model the evidence is not all
that clear since the model version with nonseparable utility still leaves the employment fluctuation somewhat smaller than what postwar U.S. data show. In this paper, I present results suggesting that this could be due to measurement errors in the data relative to the model, in particular due to the fact that data on hours are not measured in efficiency units. Further work, perhaps building on that model, might allow for two types of labor skills. This could formally be done in at least a couple of ways. One could allow preferences (and also the time allocation) of the two types of individuals to differ. Hours measured in efficiency units could be perfectly substitutable in production, and one could characterize the equilibria by solving the stand-in problem in which utility is a weighted sum of the two utility functions. An alternative would be to emphasize less the differences in utility, but instead allow for trade-offs in production combined with differences in time allocation for market work (perhaps not unreasonable for a representative household).

In trying to reconcile the differences in supply elasticities that have been estimated in various studies, it is important to consider the possibility that some empirical models may have been set up in such a way that there is a limit to how large the elasticity can be. The results in Section 3 suggest that if one estimates the parameters of a utility function which is assumed from the outset to be time-separable in leisure and it is appropriately restricted to account for a priori information from long-run behavior and from other applied areas, then a short-run elasticity of 1.5 is about the most one could expect to find regardless of what is in the data. Also, the importance of allowing for (perceived) permanent wage changes to have little or no effect on labor supplied was emphasized.
A two-parameter specification of the effect of past leisure choices on current utility was shown to be consistent with a household-production theory in which nonmarket time contributes to building up or maintaining the stock of a generally unobservable durable whose services are an argument of an inherently time-separable utility function. This theory also suggests that due to intertemporal aggregation it would not be surprising to see data sets with different lengths of the observation period yield very different results.

Finally, the possibility of getting information from panel data was discussed. Because of severe measurement problems, perhaps combined with the relatively small size of the reasonably reliable part of the subsample on wives, some pessimism is expressed in this regard. Or, to put it differently, it will most likely prove to be an extraordinarily challenging empirical task to get reliable information on the importance of nonseparability from the panel data.
Footnotes

1. See Lucas (1975), Barro (1976), and subsequent work.

2. For a recent statement on the need for this approach, see Sargent (1982).
   See also, in particular, Lucas (1976).

   See also Black (1979).


5. I am here referring to the infinitely lived agent framework, which appears particularly convenient for being confronted with data, including the imposition of a priori restrictions. An alternative optimizing framework, of course, is the overlapping-generations model, which may be more appropriate for theoretical analysis when one is not yet serious about the data. See McCallum (1982) for an evaluation of its usefulness.

6. See Hodrick and Prescott (1981), who used U.S. data from 1951:1 to 1979:2. In this paper, I shall refer to statistics using the same measure of cyclical fluctuation, but with the sample period starting and ending three years later.

7. In the model, the equilibrium real wage rate is proportional to productivity and for that reason displays exactly the same percentage fluctuation as productivity.

8. Part of the reason for the high correlation in the model between output and productivity is undoubtedly the way in which technological shocks enter multiplicatively in the production function.
9. Kydland and Prescott (1982) mention the potential importance of these types of measurement errors, but do not actually use them in their analysis or report on their quantitative implications.

10. Kennan (1983) presents results which he says "indicate that it will be difficult to distinguish between movements along a nonseparable supply function, on the one hand, and shifts in a separable supply function, on the other." This may be true within the context of a particular model, but these will probably be among the easier alternatives to find ways of distinguishing between. At this stage of development of business cycle theory, it is probably best to be careful not to bring too much into the theory through the stochastic structure. That would tend to add too many free parameters whose realism is hard to judge. A way to interpret the nonseparable utility function is that it puts structure on the way in which the schedule between current hours supplied and the real wage shifts over time.

11. Without loss of generality, we assume that the total per-period time allocation is one, so that \( n_t = 1 - l_t \).

12. See Kydland and Prescott (1982, pp. 1355-56) for a derivation using a similar utility function, although in a different context. In the present case, long-run consumption, \( c \), is \( w_n \).

13. For the purpose of determining the variance-covariance properties, the two specifications were linearized in the following sense. Consumption was substituted for from the budget constraint into the utility function and second-order Taylor series approximations were made around the long-run values. Decision rules for hours worked and net assets were then determined recursively.
14. The timing decision itself may tie in here in an interesting way, although I do not consider that particular aspect here. My colleague, Joe Hotz, is currently working on that problem.

15. Eichenbaum, Hansen, and Singleton (1982), using monthly aggregate data, allow both lagged leisure and consumption to enter the current utility function.

16. See also Bernanke (1982). Accounting for expenditures on durables and the services from the stock could be important for interpreting consumption studies such as Hall (1978) and Sargent (1978), in particular the findings in relation to the permanent-income hypothesis.

17. See, e.g., Heckman and MaCurdy (1980), MaCurdy (1981), and Hotz, Kydland, and Sędłacek (1982).

18. In the Kydland-Prescott model, elements of the technology served much of that function.


20. Some researchers have devoted a great deal of attention to the nonlinear budget constraint. See, for example, Johnson and Pencavel (1983) who estimate labor supply relations using panel data from the Seattle and Denver Income Maintenance Experiment. They employ a specification in which hours lagged one year enter the supply function, but find rather small elasticities.


22. I am grateful to Tom MaCurdy and Tom Mroz for making this sample available to me.
23. In the basic nonseparable example where these correlations, or in this case their means, were -0.28 and -0.36 as reported in Section 3, their standard deviations were 0.23 and 0.22, respectively. For the purpose of computing these standard deviations, each individual in the model was assumed to start off at his long-run value of hours and zero net assets, and 18-period samples were created, of which the last 12 periods were used in computing the statistics for each individual. The standard deviations are based on 1000 individuals.


References


